

Neurálne modely v kognitívnej robotike: porozumenie a pomenovávanie akcií

Kristína Rebrová, Igor Farkaš

Centrum pre kognitívnu vedu, KAI FMFI UK, Univerzita Komenského
Mlynská dolina, 84248 Bratislava
E-mail: {rebrova,farkas}@fmph.uniba.sk

Abstrakt

Jednou z ústredných tém v súčasnej kognitívnej vede je problematika ukotvenia jazyka. Nedávne objavy v neuro-psychológii a kognitívnej lingvistike poukazujú na to, že jazyk nemožno oddeliť od percepcie a akcie. Naopak, porozumenie jazyku sa viaže na multimodálnu simuláciu prebiehajúcu v tých istých častiach mozgu, ktoré zodpovedajú za priamu percepciu a akciu. Tento fenomén bol zachytený v množstve výpočtových modelov používaných na riadenie fyzických a simulovaných robotov. V príspevku predstavíme vybrané modely prepojenia akvizície jazyka s percepciou a motorikou založené na umelých neurónových sieťach. Načrtne špecifiká, ktoré by mal zohľadniť náš model akvizície jazyka implementujúci porozumenie zjednodušenému jazyku pre akcie vykonávané rukou na základe teórie motorickej rezonancie a zrkadliacich neurónov.

1 Úvod

Kognitívna robotika je v súčasnosti jednou z najdynamickejších a najpopulárnejších oblastí v poli kognitívnych vied. Jej základom je tzv. konštruktivistický prístup (Pfeifer a Scheier, 1999, pozri aj Takáč, 2009), ktorého ústrednou myšlienkou je skúmanie rôznych kognitívnych procesov pomocou ich modelovania vo fyzických a simulovaných robotoch. Takýto prístup sa, na rozdiel od klasických dizajnerských postupov, ktoré pri vytváraní inteligencie postupujú výhradne zhora (od dizajnéra k robotom), zaoberá inteligenciou a schopnosťami agentov im vlastnými, ekologicky dôležitými, a dáva dôraz na ich emergenciu a učenie sa. Ďalším dôležitým faktorom v tejto oblasti je vtelenosť kognície. Vtelený alebo stelesnený agent má fyzické telo so senzormi a efektormi, ktoré mu umožňuje interagovať s prostredím, v ktorom sa nachádza. V prípade vtelenej kognície sa teda na kognitívnych schopnostiach agenta, vrátane usudzovania či jazyka, podieľa celé jeho telo a tiež prostredie, ktoré nie je oddeliteľné od agenta a jeho inteligencie (Pfeifer a Scheier, 1999). Vzhľadom na to, že skonštruovať agenta, ktorý by bol kognitívne na úrovni či už dospelého ale aj mladého človeka je úloha tak komplexná a zložitá, sa osvedčili prístupy, pri ktorých sa

študujú a modelujú jednotlivé kognitívne kapacity oddelene, a najmä, v zjednodušených scenároch a prostrediach (mikrosvetoch).

Dôležitým aspektom kognitívnej robotiky je aj jej prepojenie s aktuálnymi poznatkami z psychológie a neurovedy, ktoré môžu byť implementované v (humanoidných) robotoch. Okrem vtelenej kognície je dôležitou paradigmou v kognitívnej robotike, ako aj v rôznych ďalších kognitívnych disciplínach, teória ukotvenej kognície (grounded cognition). Kľúčovým podnetom pre túto oblasť bola teoretická námietka Harnarda (1990) voči dizajnerským postupom. Aby bol systém ekologicky validný a inteligentný v pravom slova zmysle, musia byť významy, vnútorné reprezentácie všetkého s čím sa agent stretne, vlastné agentovi a nie človeku, ktorý ho programoval. Podobne ako Searle (1980) v čínskej izbe, ktorý sám nerozumie po čínsky, by ani umelé agenty nemali „len“ manipulovať s arbitrárnymi amodálnymi symbolmi, vznikajúcimi zo zmyslových podnetov pomocou neznámeho procesu transdukcie (Fodor, 1975), a po spracovaní ich rovnakým bližšie nešpecifikovaným procesom premieňať na svoje akcie. Paradigma kognitívizmu, inšpirovaná novými objavmi v matematickej logike, lingvistike a informatike, v centre ktorej sú amodálne symboly (Fodor, 1975) a univerzálna gramatika (Chomsky, 1965), dominovala v kognitívnej vede zhruba od polovice minulého storočia. Arbitrárne amodálne symboly si môžeme predstaviť ako zoznamy príznakov, sémantické siete, rámce, schémy, či vety predikátovej logiky.

Jednou z najvplyvnejších reakcií na túto paradigmu bola teória perceptuálnych symbolov (Barsalou, 1999). Na rozdiel od amodálnych symbolov, ktoré sa nepodobajú na vnemy, z ktorých vzišli, perceptuálne symboly sú podmnožinami perceptuálnych stavov v rámci senzomotorických systémov, ich vnútorná štruktúra je modálna a podobá sa na príslušné perceptuálne stavy. Ako ukázali moderné disciplíny kognitívnej vedy, ako neuropsychológia či kognitívna lingvistika, vyššie kognitívne schopnosti, vrátane usudzovania či jazyka, sú ukotvené v percepcii a motorike, interakcii agenta s vlastným telom a prostredím a tiež v sociálnej interakcii medzi agentmi.

V príspevku, ktorý voľne nadväzuje na (Farkaš, 2010) načrtujeme základ empirickej podpory pre ukotvené teórie porozumenia akciám a jazyku. Ďalej predstavíme modely akvizície zjednodušeného jazyka popisujúceho akcie, ktoré priamo prepájajú jazyk s percepciou a motorikou, založené na umelých neurónových sieťach osadených do fyzických robotov. Nakoniec popíšeme špecifikáciu budúceho výskumného projektu, modelu porozumenia akciám prepojeného s akvizíciou jednoduchého jazyka, ktorý tieto akcie pomenováva.

2 Empirická podpora ukotveného porozumenia

Empirické štúdie podporujúce simulačné teórie porozumenia predpokladajú, že na porozumení pozorovaným akciám alebo ich verbálnym popisom sa priamo podieľajú oblasti mozgu, ktoré sú zodpovedné za produkciu týchto akcií, resp. sekvencií pohybov. Podobne ako v Barsalouovej teórii perceptuálnych symbolov, ktorá predpokladá porozumenie na báze multimodálnej perceptuálnej simulácie, simulácia z pohľadu kognitívnej neurovedy zahŕňa nepriamu aktiváciu tých istých neurálnych okruhov ako pri aktívnej percepcii a akcii (Borghini a spol., 2010). Na základe tohto vteleného, automatického procesu sme schopní porozumieť správaniu a úmyslom iných (Gallese, 2009).

2.1 Porozumenie akciám a zrkadliace neuróny

Aby sme sa dostali k jadrú akvizície jazyka, potrebujeme poznať fungovanie elementárnejšieho porozumenia akciám (action understanding). Pod týmto pojmom rozumieme v prvom rade schopnosť agenta (človeka či zvierat'a) pozorujúceho ďalšieho agenta vykonávajúceho nejakú akciu (napr. uchopenie kúska jedla a vloženie si ho do úst) rozlíšiť, že pozorovaný agent niečo vykonáva. Súčasne musí agent rozpoznať vykonávanú akciu od iných akcií a vedieť vhodne zareagovať (Gallese a spol., 1996). Táto posledná časť definície môže znamenať, že agent urobí akciu, ktorá je preňho prospešná, prípadne, keď prejdeme z prírody do sveta experimentov, akciu pomenuje alebo inak naznačí, že jej rozumie. Porozumenie akciám nutne zahŕňa vizuálne vnímanie, otázne však je, či pre úplné porozumenie stačí len analýza vizuálnych vlastností, alebo je potrebný ďalší mechanizmus, ktorý nám napomáha „preveteliť sa“ do pozorovaného aktéra.

Rizzolatti a spol. (2001) rozlišujú dve základné hypotézy o porozumení akciám. Prvá, vizuálna hypotéza funguje čisto na báze vizuálnej analýzy jednotlivých častí pozorovanej scény. Táto hypotéza je bližšie ku klasickým, symbolovým teóriám kognície. Naopak,

hypotéza priameho párovania, ktorá je príkladom simulačnej teórie kognície, tvrdí, že pre porozumenie akcií je potrebné akciu vnútorne pripodobniť k niečomu, čo pozorujúci agent už pozná. Na tento účel slúži podľa autorov teórie *motorická rezonancia*, ku ktorej dochádza v *zrkadliacich neurónoch*.

Zrkadliace neuróny boli pôvodne objavené u opíc druhu *Macaca nemestrina* (Pellegrino a spol., 1992) v premotorickej oblasti F5, ktorá je aktívna, keď opica vykonáva rôzne typy uchopení a iných akcií rukami a ústami. Sú to subpopulácie neurónov tejto oblasti, ktoré sú aktívne nielen vtedy, keď opica akciu vykonáva, ale aj vtedy, keď tú istú (podobnú) akciu pozoruje. Po dlhšej dobe dohadov boli zrkadliace neuróny objavené pomocou priamych meraní aktivity neurónov (single-cell recording) aj u ľudí (Mukamel a spol., 2010). Na čiastočnú aktiváciu motorických oblastí, ich rezonanciu, spôsobenú čisto zrakovým vnemom sa môžeme pozeráť ako na motorickú simuláciu pozorovaného deja, ktorá môže slúžiť na jeho rozpoznanie a vyvolanie správnej reakcie naň. Viac o zrkadliacich neurónoch viď v článku Farkaš a spol. v tomto zborníku.

Spolu so zrkadliacimi neurónmi obsahuje oblasť F5 aj takzvané kanonické neuróny. Tie tiež reagujú na vizuálne podnety, no len v prípade, že sa na scéne nachádza objekt, s ktorým sa dá potenciálne manipulovať, ale nie aktér, čo niečo vykonáva. Kanonické neuróny reprezentujú afordancie objektu (Gibson, 1977). Ich aktivácia nereprezentuje to, čo sa deje, ale svojím spôsobom simulujú to, čo by mohol pozorovateľ s objektom urobiť.

Rizzolatti a Arbib (1998; Arbib, 2005) postulovali, že zrkadliace neuróny zohrávajú úlohu nielen pri porozumení akciám, ale aj ako „chýbajúce prepojenie“ medzi komunikáciou zvierat a ľudským jazykom. Autori vychádzajú z toho, že oblasť F5 je potenciálnym homológom¹ Brocovej oblasti, ktorá sa zapája pri produkcii reči, ale aj pri porozumení komunikačným gestám. Podľa tejto teórie zabezpečila schopnosť zrkadliacich neurónov asociovať produkciu akcií s ich vnímaním, a teda rozšíriť porozumenie medzi komunikujúcimi jedincami, evolučnú výhodu pre vznik jazyka u ľudí.

2.2 Porozumenie jazyku a motorická rezonancia

Podobne ako pri vnímaní a porozumení akciám, dochádza k aktiváciám motorických oblastí aj pri spracovávaní viet či slov, ktoré akcie pomenúvajú. Pulvermüller a kolegovia (2001; tiež Pulvermüller, 2005; Hauk a spol., 2004) ukázali, že čítanie či počúvanie slovies pre rôzne pohyby vykonávané rôznymi časťami tela (napr. zdvihnúť/ruka,

¹ oblasťou so spoločným predchodcom

lízat/ústa, kopat/ noha) vyvolalo somatotopickú aktiváciu² v motorickej kôre len 250 ms po zaznení slova. To znamená, že k motorickej rezonancii dochádza spontánne, ešte pred vedomým spracovaním podnetu, teda je vysoko nepravdepodobné, že by pri procese porozumenia malo najprv dôjsť k transdukcii a následne k aktivácii motorickej kôry amodálnymi procesmi. Fakt, že k motorickej aktivácii dochádza ihneď, spontánne a somatotopickým spôsobom napovedá, že ide o súčasť procesu porozumenia, nie o akýsi jeho vedľajší produkt.

Ďalším zdrojom podpory pre vtelené teórie spracovania jazyka sú výsledky rôznych behaviorálnych experimentov. Glenberg a Kaschak (2002) objavili, že pri čítaní „prenosových“ viet dochádza k *efektu kompatibility akcie a vety* (action-sentence compatibility effect, ACE). Tento jav demonštruje, že sa pri porozumení vetám zapája motorika na základe princípu interferencie. V prípade, že sa smer prenosovej vety nezhoduje so spôsobom, ako treba na ňu reagovať, je reakčný čas vyšší ako v prípade zhody. Ak mal napríklad proba rozhodnúť o tom, či je veta „Janka ti podala knihu“ zmysluplná alebo nezmyselná, tým, že stlačí tlačidlo ktoré je ďalej od jeho tela (áno) oproti tlačidlu pre opačnú odpoveď (nie), bude reakčný čas dlhší ako v prípade, keby mal rukou robiť pohyb kongruentný so smerom, ktorý implikuje veta, teda „od Janky ku mne“ (tlačidlo „áno“ by muselo byť bližšie pri probandovi). Neskôr Glenberg a spol. (2008) ukázali, že tento jav sa vzťahuje nielen na konkrétne vety (podávanie knihy), ale aj na abstraktné vety (napr. odovzdanie myšlienky).

3 Výpočtové modely porozumenia akciám a akvizície jazyka

V tejto časti sa budeme venovať vybraným výpočtovým modelom ukotvenej akvizície jazyka o akciách na báze neurónových sietí. Tieto modely poukazujú na rôzne aspekty prepojenia percepcie, akcie a jazykových inštrukcií, zapájajú rôzne paradigmy učenia a neurálne architektúry. Ich spoločným znakom je, že dávajú dôraz na neoddeliteľnosť jazyka od percepcie a akcie, a tiež ukazujú, ako by mohli byť tieto aspekty kognície prepojené pomocou učenia. Všetky modely, ktoré uvedieme boli osadené do fyzických alebo simulovaných robotov, a teda sú aj ich charakteristiky a funkčnosť závislé od ich senzorov a efektorov.

3.1 Model prenosu ukotvenia

² Primárna motorická a premotorická kôra sú organizované somatotopicky. Podobne ako na mape, rôzne lokality kôry reprezentujú rôzne časti tela.

Zaujímavým aspektom porozumenia na základe ukotvenia symbolov (slov) je otázka prenosu významu z dvoch ukotvených komponentov, napríklad „kôň“ a „prúžky“, do nového konceptu, napr. „zebra“, ktorý bude v týchto dvoch konceptoch ukotvený nepriamo. Cangelosi a Riga (2006) modelovali *prenos ukotvenia* v simulovanom robotickom agentovi riadenom doprednou neurónovou sieťou. Tento robotický agent je vybavený telom s 3-segmentovými rukami so 6 stupňami voľnosti a žije v prostredí vytvorenom pomocou Open Dynamics Engine (ODE, www.ode.org). Agent sa učí na základe imitácie druhého agenta, učiteľa, ktorý je vopred predprogramovaný, aby predvádzal rôzne akcie sprevádzané ich jazykovým popisom – pomenovaním. Keď sa agent naučí vykonávať a pomenovávať základné akcie je schopný sa autonómne učiť akcie vyššej úrovne, len na základe jazykového popisu zadaného experimentátorom. Napríklad v príkaze „chytiť (je) zovrieť-l'avú ruku (a) zovrieť-pravú-ruku“ sa agent učí akciu vyššej úrovne „chytiť“ ako kompozíciu dvoch základných akcií, ktoré pozná.

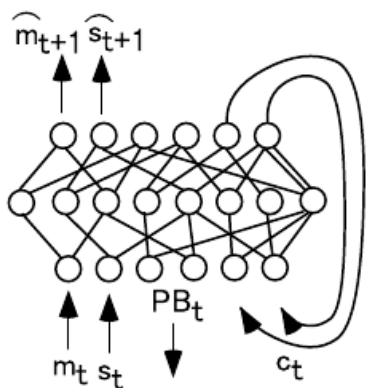
Architektúra riadiacej neurónovej siete je relatívne jednoduchá. Na vstupnej vrstve agent dostáva informáciu z rôznych modalít, konkrétne vizuálnu (virtuálna sieťnica), jazykovú a motorickú (natočenia kĺbov a sila v motoroch). Ďalej má sieť skrytú vrstvu a výstupnú vrstvu, na ktorej je opäť informácia rôznych druhov – motorická a jazyková. Takto môže agent prijímať jazykové inštrukcie ako aj pomenovávať akcie, ktoré vykonáva bez jazykového vstupu. Pre učenie zvolili autori štandardný algoritmus spätného šírenia chyby (error back-propagation). V procese učenia si agent najprv „odporozuje“ a vykoná pohyby demonštrátora, ktoré sú porovnané so želaným výstupom vytvoreným pomocou špeciálnej funkcie vyrobí odhad uhlov kĺbov a síl aplikovaných na motory. Podobne jazykový vstup je spracovaný a lokalisticky zakódovaný do aktivácií na vstupe a porovnaný s aktiváciou na výstupe. Chyba, ktorá vznikne porovnaním želaného a reálneho výstupu je následne spätne šírená sieťou a premieta sa do zmien váh siete. Napriek tomu, že algoritmus spätného šírenia chyby vedie k úspešnému naučeniu sa základných aj odvodených akcií a pomenovaní, je všeobecne považovaný za biologicky neplauzibilný a bolo by vhodné ho nahradiť iným, plauzibilnejším mechanizmom, napríklad učením s posilňovaním (viď nižšie).

3.2 Model RNNPB

Rekurentná neurónová sieť s parametrickými biasmi (RNNPB) (Tani, 2003; Tani a Ito, 2003) je pozoruhodná neurálna architektúra, ktorá je schopná meniť svoje dynamické správanie na základe vopred natrénovaných vzorov uložených do pamäte pomocou špeciálnych

vektorov parametrických biasov. Tie si počas tréningu vytvára sieť sama, samoorganizáciou, čiže sú vlastné samotnej sieti, nie zadané programátorom. Ten ich však môže po natréningu zo siete vyabstrahovať a následne ich zadávať na vstup a tak určovať správanie siete.

Modelová architektúra siete je zobrazená na obrázku 1. RNNPB je rekurentná sieť Jordanovho typu (Jordan, 1990). Parametrické biasy sú v sieti kódované v podmnožine neurónov vstupnej vrstvy. Tá okrem vstupu z rôznych modalít obsahuje aj kontextové neuróny, ktoré prijímajú signál z výstupných neurónov, informáciu z predošlého časového úseku. Ďalej má sieť skrytú a výstupnú vrstvu. V bežnom prípade správania sa pôsobí sieť ako dopredný model (Kawato a spol., 1987), ktorý predpovedá nasledujúci stav (agenta a prostredia) zo stavu súčasného. Celková matica váh v sieti ako aj sada PB vektorov sa trénuje pomocou spätného šírenia chyby v čase (BPTT) (Rumelhart a spol., 1986). Po naučení je sieť schopná generovať sekvencie správania na základe zadaných vstupov ako aj vopred vložených PB vektorov, ktoré môžu pochádzať z tréningovej fázy siete alebo z inej RNNPB, ako uvidíme v nasledujúcom príklade. Podľa Taniho a kolegov (2004) sú PB vektory funkčným ekvivalentom zrkadliacich neurónov, pretože multimodálne spájajú vstupy, sensorické a motorické, a s nimi asociované správanie.



Obr. 1. Schematický náčrt RNNPB

V konkrétnom prípade implementácie RNNPB do robotického agenta učiaceho sa jazyk, Sugita a Tani (2005) rozšírili architektúru na dve slabšie prepojené RNNPB siete. Ich prepojenie spočíva v simultánnom učení pomocou *previazania* PB neurónov. Cieľom tohto špeciálneho učenia je dosiahnuť zhodu, v ideálnom prípade identitu, medzi sadami PB vektorov týchto dvoch sietí. V kontexte kognitívnej robotiky reprezentujú tieto dve siete dve modality, či skôr dva aspekty akvizície jazyka, a to percepciu a motoriku v jednej sieti a jazykové schopnosti v sieti druhej. Ako zdôrazňujú sami autori, napriek tomu, že je jazyková sieť schopná naučiť sa

správne syntax a kombinatorickú štruktúru jednoduchého jazyka (vety typu sloveso + podstatné meno), bez prepojenia so sieťou zabezpečujúcou percepciu a akciu, nebude agent samotnému jazyku rozumieť. Preto možno v tomto prípade hovoriť o vtelenom systéme porozumenia len pri plnej funkcionalite oboch sietí.

V experimentoch bol použitý mobilný robot, ktorý mal dve kolesá, ruku s jedným kĺbom a farebnú kameru, pričom jeho úlohou bolo interagovať s tromi objektami v mikrosvete odlišenými tromi farbami na stálych pozíciách, kvôli skúmaniu schopnosti zachytiť kompozicionalitu jazyka. Výsledky experimentov ukázali, že agent, ktorý sa trénuje súčasne vykonávať akcie a aktivovať jazykové reprezentácie, je následne schopný správne reprodukovať tieto akcie podľa jazykových pokynov aj v prípade, že ide o nové kombinácie slovies, teda činností a objektov, ktoré nikdy nezažil.

V ďalšom experimente (Sugita a Tani, 2008) sa autori zamerali na učenie sa akcií zameraných na cieľ a transfer skúsenosti do systematických znalostí pomocou simulovaného mobilného robota podobného tomu z predošlého experimentu. V tomto prípade je architektúra RNNPB mierne pozmenená a skladá sa z jednej doprednej 4-vrstvovej siete pre percepciu a akciu, ktorá je spojená s ďalšou sieťou pre jazyk a PB vektory pomocou spojení druhého rádu do výstupnej vrstvy prvej siete. Druhá sieť teda pôsobí ako ovládací prvok pre doprednú sieť riadiacu robota, ktorý je schopný vykonávať akcie na základe vnemov z prostredia ako aj na základe jazykových inštrukcií. Výsledky rôznych experimentov s týmto modelom ukázali, že agent bol schopný generalizovať na základe dvoch schopností, prenášania schopnosti do nového prostredia (s novými objektami) a kombinovania naučených akcií do nových akcií. Analýza klasterizácie hlavných komponentov PB vektorov ukázala, že agent úspešne generalizoval znalosti z predložených viet a vytváral si znovu použiteľné koncepty, podobne ako to robia malé deti (Tomasello, 2003).

Napriek tomu, že architektúra RNNPB je zaujímavá a inováčná z hľadiska funkcionality založenej na schopnosti meniť svoju dynamiku, čo je samo o sebe biologicky prijateľné, nachádza sa v nej hneď niekoľko problémov. Hlavný z nich spočíva v metóde učenia, keďže BPTT nie je ekologicky validná metóda. Jej ďalším nedostatkom je, že je náročná na čas a výpočtové prostriedky, na ktorých je implementovaná, čo robí z tréningu RNNPB náročný proces, neporovnateľný s ľahkosťou, s akou získavajú zručnosti malé deti. Slabou stránkou modelu je aj spôsob učenia sekvencií: model využíva požadované proprioceptívne signály v každom kroku (ak keby matka učila dieťa tak, že mu držala ruku a ťahala ju požadovaným smerom). Na druhej strane je nutné dodať, že tento model a jeho aplikácie, sú na

rozdiel od mnohých iných skutočne vtelené a ukotvené v sensoricko-motorickej interakcii agenta s prostredím.

3.3 Model priameho ukotvenia slov v akcii

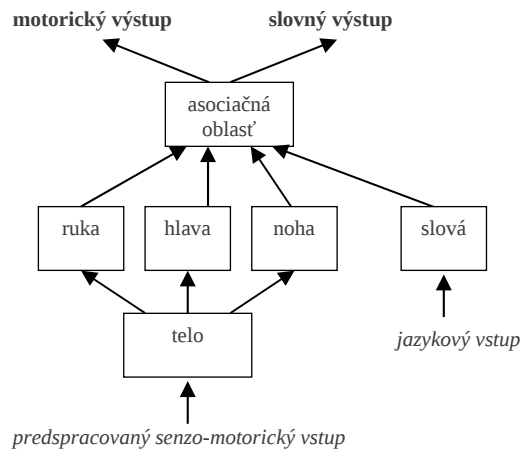
V nedávno publikovanej štúdií, Marocco a spol. (2010) modelovali akvizíciu slov pre akcie v humanoídnom robotovi iCub (Tikhanoff a spol., 2008), respektíve jeho simulátore (kvôli zjednodušeniu a technickým obmedzeniam). V tomto prípade sa využíva iba podmnožina zo sensorov a efektorov robota a nový parameter, ktorému autori hovoria „okružnosť“, vypočítaný z obrazu, ktorý robot vidí. Práve tento parameter sa asocjuje s agentovými možnosťami interagovať s objektami a s jazykovým vstupom. Robot prichádza do kontaktu vždy len s jedným objektom, tým istým spôsobom, a to tak, že sa ho snaží potlačiť do jednej strany a sleduje, čo sa s objektom stane, akú silu na to musí využiť a či sa objekt pohol. Podobne ako v predošlej štúdií bola na riadenie robota bola použitá rekurentná neurónová sieť s rovnakým počtom a „modalitou“ neurónov na vstupnej a výstupnej vrstve tvoriaca dopredný model. Na trénovanie bol aj tu použitý algoritmus BPTT (kritiku vid' vyššie).

Zaujímavým aspektom tohto modelu a výsledkov experimentov s ním je, že sa agent naučil slová (lokalisticky kódované) pre spôsoby manipulácie s predmetmi, ako „gúľajúci sa“, „súchajúci sa“ a „nehybný“ (objekt). Ide tu o priame ukotvenie arbitrárnych kódov v percepcii a akcií. Agent v podstate nerozlišuje medzi tvarmi či farbami objektov, ale pomenováva len ich vlastnosti podľa jeho možnej interakcie s nimi bez akejkoľvek dodatočnej informácie od experimentátora.

3.4 Projekt MirrorBot

Model samoorganizovanej pamäte pre akcie a slová (Wermter a Elshaw, 2003) bol navrhnutý v rámci projektu MirrorBot, zameraného na emergenciu reprezentácií akcií, vnemov a konceptov vo fyzickom robotovi vybavenom zrkadliacimi neurónmi. Štúdia inšpirovaná objavmi Pulvermüllera a kolegov (2001) o somatotopickej aktivácii motorických častí mozgu jazykovými podnetmi, ponúka štruktúrovaný model na báze samoorganizujúcich sa máp (SOM) (Kohonen, 1997). Model, schematicky zobrazený na obrázku 2., pozostáva zo vstupnej mapy pre vnemy, ktorá dostáva na vstup vopred získané a predspracované dáta zo sensorov robota. Z nej sa šíri aktivácia do 3 ďalších sietí, z ktorých každá zastupuje jednu časť tela, hlavu, ruku a nohu. Paralelne s touto sensorickou časťou funguje mapa, ktorá reprezentuje slová, ktoré sa agent učí. Tieto dva moduly

sú napojené na asociačnú oblasť (tiež SOM), ktorá produkuje výstup buď to vo forme akcie robota alebo slovného pomenovania.



Obr. 2. Schematický náčrt modelu samoorganizovanej pamäte

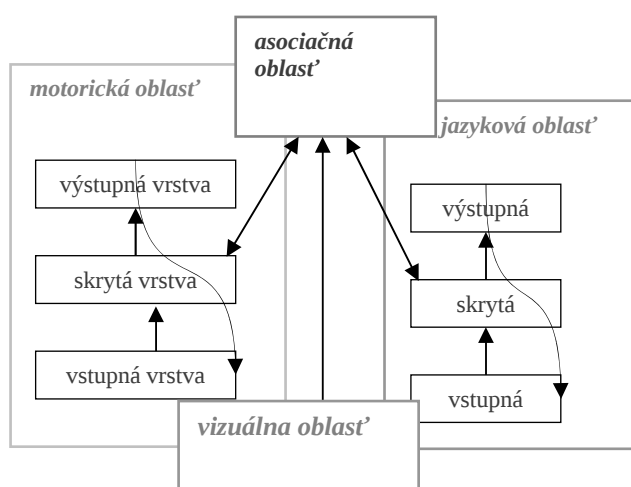
Asociačná oblasť zohráva podobnú úlohu ako naviazané PB vektory v dvoj-sieťovom RNNPB modeli, avšak zo špecifikácie v spomínanom článku nie je jasné, akým spôsobom sa spracúva výstup z asociačnej oblasti do podoby konkrétnych výstupov. Naopak v RNNPB architektúre sa na výstupnej vrstve vždy priamo generujú pokyny pre efektor robota. Čo sa týka výsledkov experimentov popísaných v (Wermter a Elshaw, 2003), autori overovali len fungovanie 4 „sensorických“ sietí, preto nie je možné tvoriť celkové závery o funkčnosti, použiteľnosti, či škálovateľnosti tohto modelu. Pri vyhodnocovaní jeho kvalít je nutné vyzdvihnúť jeho relatívne vysokú biologickú prijateľnosť, keďže ide o model s Hebbovským učením fungujúcim aj v mozgu, a bez učiteľa, inšpirovaný objavmi z modernej neurovedy. Medzi nedostatky tohto modelu patrí, že na rozdiel RNNPB pracuje len ako asociátor bez zahrnutia dynamických vlastností do systému (t.j. schopnosti predikovať nasledujúce senzomotorické hodnoty).³ Na druhej strane, peknou vlastnosťou tohto modelu je topografická organizácia neurónov (vlastnosť mozgovej kôry).

4 Náš model

V tejto časti príspevku popíšeme vlastnosti modelu, ktorý budeme vytvárať v rámci dizertačnej práce K. Rebrovej „Modelovanie akvizície jazyka pomocou neurónových sietí“. Ide o model akvizície zjednodušeného jazyka

³ Akcie boli síce kódované ako sekvencie vzorov, no boli vstupným modulom (SOM pre telo) spracovávané ako priestorové vzory.

u simulovaného robotického agenta (iCub simulátor) založený na umelých neurónových sieťach. Agent sa bude počas svojho života postupne učiť vykonávať jednoduché akcie rukami, ako natáhanie sa, chytanie predmetu, zhodenie predmetu a podobne a tiež porozumieť týmto akciám na základe motorickej rezonancie pri ich pozorovaní. Neskôr sa k učeniu akcií pridá aj učenie sa jazyka – jednoduchých slov pomenovávajúcich objekty a akcie, ktoré s nimi vie vykonať v podobe jednoduchých, niekoľko-slovných viet. Počas učenia by mal agent dostávať informáciu o úspešnosti svojho konania a tiež jazykové vstupy od druhého agenta v roli matky – učiteľky. V želanom výslednom stave by mal byť agent schopný pomenovávať nie len akcie, ktoré vykonáva on sám, ako vo vyššie uvedených modeloch, ale aj akcie, ktoré pozoruje.



Obr. 3. Schematický náčrt nášho modelu

Riadiaca architektúra agenta, načrtnutá na obrázku 3., bude pozostávať z niekoľkých podsietí, líšiacich sa či už funkciou ale aj typom architektúry. Samotné pohyby agenta by mali byť riadené rekurentnou neurónovou sieťou, na ktorej výstupe by mali byť neuróny zodpovedajúce priamo výsledným, dopredným modelom predpovedaným hodnotám agentových efektorov – uhlov kĺbov, síl motorov a podobne. Na vstupe bude figurovať predspracovaný vizuálny vstup, verbálny vstup a motorická aktivácia z predošlého kroku. Súčasťou modelu by mali byť aj vyššie asociačné oblasti, napríklad vyššia motorická oblasť, ktorá bude súčasne dostávať vizuálne a motorické vstupy z vykonávaných akcií a na základe princípu širokej kongruencie, podobne ako u skutočných zrkadliacich neurónov (Rizzolatti a Sinigaglia, 2010), (invariantne) reagovať na celé kategórie pohybov počas vykonávania celej sekvencie. Tieto neuróny, natréňované súčasne na motorických a vizuálnych vstupoch budú reagovať aj v prípade

neprítomnosti motorického stimulu v duchu motorickej rezonancie. Ďalšou časťou bude sieť, ktorá bude spracovávať verbálne vstupy, bude prepojená s vyššími oblasťami a v súlade so súčasnými empirickými poznatkami bude tiež schopná spôsobovať motorickú rezonanciu a teda porozumenie verbálne opísanej akcii. Motorická rezonancia bude slúžiť nie len na generovanie správnej jazykovej, ale aj správnej motorickej odozvy, napríklad v prípade imitácie.

Čo sa týka použitia paradigiem učenia neurónových sietí, chceli by sme sa zamerať na biologicky prijateľnejšie metódy učenia ako je samoorganizácia (napríklad v asociačných oblastiach ako u Wernter a Elshaw, 2003) a učenie s posilňovaním (reinforcement learning, RL). Pri RL dostáva sieť informáciu v podobe skalárnej veličiny (odmeny alebo trestu) o tom, aké „dobré“ sú akcie vykonávané agentom (Doya, 2000). Túto informáciu, ktorá nepredstavuje priamy chybový signál dostáva sieť ako celok, nie jej jednotlivé neuróny. Ide o biologicky prijateľnejšiu formu učenia, veď ani malé dieťa nedostane pri učení sa uchopenia predmetov šnúrku na ruku, čo by ho presne viedla. Problematickou oblasťou tohto prístupu je časová zložitosť, keďže pri menej jasnej, nepriamej odozve môže sieť potrebovať viac iterácií na konvergenciu.

5 Záver

Problematika ukotvenia jazyka pre akcie je jednou z aktuálnych tém na poli kognitívnej robotiky. Nedávne objavy v neuropsychológii a kognitívnej lingvistike poukazujú na to, že jazyk nemožno oddeliť od percepcie a akcie. Vtelené (embodied) porozumenie jazyku predpokladá simuláciu na strane pozorovateľa, pri ktorej dochádza ku motorickej rezonancii – aktivácii tých istých neurálnych okruhov ako pri aktívnej percepcii a akcii. V príspevku sme načrtli vybrané modely akvizície jednoduchého jazyka pre pomenovávanie akcií na báze neurónových sietí osadené do vtelených agentov, ich architektúry, metódy učenia, ich výhody aj nedostatky. V závere sme predstavili teoretický návrh nášho modelu, ktorý sa od predošlých príkladov líši práve v dôraze na motorickú rezonanciu a jej dôležitej úlohe pri oddelení pomenovávania akcie od jej priameho vykonávania.

Pod'akovanie: Tento príspevok vznikol v Centre pre kognitívnu vedu na Katedre aplikovanej informatiky, FMFI UK v Bratislave, za podpory grantovej agentúry VEGA MŠ SR, v rámci grantových úloh 1/0439/11 a 1/0602/10.

Literatúra

- Arbib, M. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral a Brain Sciences* 28(02), pp. 105–124.
- Barsalou, L. (1999). Perceptual symbol systems. *Behavioral a brain sciences* 22(04), pp. 577–660.
- Borghì, A., C. Gianelli, a C. Scorolli (2010). Sentence Comprehension: Effectors a Goals, Self a Others. An Overview of Experiments a Implications for Robotics. *Frontiers in Neurorobotics* 4.
- Cangelosi, A. a T. Riga (2006). An embodied model for sensorimotor grounding a grounding transfer: Experiments with epigenetic robots. *Cognitive science* 30(4), pp. 673–689.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. The MIT press.
- Doya, K. (2000). Reinforcement learning in continuous time a space. *Neural computation* 12(1), pp. 219–245.
- Farkaš, I. (2010). Neurálne modely v kognitívnej robotike: vizuálno-motorická interakcia. *Kognice a umělý život X*, pp. 93–99.
- Fodor, J. (1975). *The Language of Thought*. Harvard University Press.
- Gallese, V. (2009). Motor abstraction: a neuroscientific account of how action goals a intentions are mapped a understood. *Psychological research* 73(4), pp. 486–498.
- Gallese, V., L. Fadiga, L. Fogassi, a G. Rizzolatti (1996, April). Action recognition in the premotor cortex. *Brain : a journal of neurology* 119 (Pt 2), pp. 593–609.
- Gibson, J. (1977). The theory of affordances. *Perceiving, acting, a knowing: Toward an ecological psychology*, pp. 67–82.
- Glenberg, A. a M. Kaschak (2002). Grounding language in action. *Psychonomic Bulletin & Review* 9(3), pp. 558.
- Glenberg, A., M. Sato, L. Cattaneo, L. Riggio, D. Palumbo, a G. Buccino (2008). Processing abstract language modulates motor system activity. *The Quarterly Journal of Experimental Psychology* 61(6), pp. 905–919.
- Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena* 42(1-3), pp. 335–346.
- Hauk, O., I. Johnsrude, a F. Pulvermüller (2004). Somatotopic representation of action words in human motor a premotor cortex. *Neuron* 41(2), pp. 301–307.
- Jordan, M. (1990). Attractor dynamics a parallelism in a connectionist sequential machine. In *Artificial neural networks*, pp. 112–127. IEEE Press.
- Kawato, M., K. Furukawa, a R. Suzuki (1987). A hierarchical neural-network model for control a learning of voluntary movement. *Biological Cybernetics* 57(3), pp. 169–185.
- Kohonen, T. (1997). *Self-organizing Maps*. Springer.
- Marocco, D., A. Cangelosi, K. Fischer, a T. Belpaeme (2010). Grounding Action Words in the Sensorimotor Interaction with the World: Experiments with a Simulated iCub Humanoid Robot. *Frontiers in Neurorobotics* 4.
- Mukamel, R., A. Ekstrom, J. Kaplan, M. Iacoboni, a I. Fried (2010). Single-neuron responses in humans during execution a observation of actions. *Current Biology*.
- Pellegrino, G., L. Fadiga, L. Fogassi, V. Gallese, a G. Rizzolatti (1992). Understanding motor events: a neurophysiological study. *Experimental brain research* 91(1), pp. 176–180.
- Pfeifer, R. a C. Scheier (1999). *Understanding intelligence*. Cambridge, MA: The MIT Press.
- Pulvermüller, F. (2005). Brain mechanisms linking language a action. *Nature Reviews Neuroscience* 6(7), pp. 576–582.
- Pulvermüller, F., M. Härle, a F. Hummel (2001). Walking or Talking? : Behavioral a Neurophysiological Correlates of Action Verb Processing* 1. *Brain a language* 78(2), pp. 143–168.
- Rizzolatti, G. a M. Arbib (1998). Language within our grasp. *Trends in neurosciences* 21(5), pp. 188–194.
- Rizzolatti, G., L. Fogassi, a V. Gallese (2001). Neurophysiological mechanisms underlying the understanding a imitation of action. *Nature Rev. Neuroscience* 2, pp. 661–670.

- Rizzolatti, G. a C. Sinigaglia (2010). The functional role of the parieto-frontal mirror circuit: interpretations a misinterpretations. *Nature reviews Neuroscience* 11(4), pp. 264–74.
- Rumelhart, D., G. Hinton, a R. Williams (1986). *Learning internal representations by error propagation*, Volume 1, pp. 318–362. Cambridge, MA: MIT Press.
- Searle, J. (1980). Minds, brains, a programs. *Behavioral a brain sciences* 3(03), pp. 417–424.
- Sugita, Y. a J. Tani (2005). Learning semantic combinatoriality from the interaction between linguistic a behavioral processes. *Adaptive Behavior* 13(1), pp. 33.
- Sugita, Y. a J. Tani (2008). A sub-symbolic process underlying the usage-based acquisition of a compositional representation: Results of robotic learning experiments of goal-directed actions. In *Development a Learning, 2008. ICDL 2008. 7th IEEE International Conference on*, pp. 127–132. IEEE.
- Takáč, M. (2009). Konštruktivistický prístup k štúdiu kognície. *Umelá inteligencia a kognitívna veda I.*, pp. 395–424.
- Tani, J. (2003). Learning to generate articulated behavior through the bottom-up a the top-down interaction processes. *Neural Networks* 16(1), pp. 11–23.
- Tani, J. a M. Ito (2003). Self-organization of behavioral primitives as multiple attractor dynamics: A robot experiment. *Systems, Man a Cybernetics, Part A: Systems a Humans, IEEE Transactions on* 33(4), pp. 481–488.
- Tani, J., M. Ito, a Y. Sugita (2004). Self-organization of distributedly represented multiple behavior schemata in a mirror system: reviews of robot experiments using RNNPB. *Neural Networks* 17(8-9), pp. 1273–1289.
- Tikhanoff, V., A. Cangelosi, P. Fitzpatrick, G. Metta, L. Natale, a F. Nori (2008). An open-source simulator for cognitive robotics research: The prototype of the icub humanoid robot simulator. In *Proceedings of the 8th Workshop on Performance Metrics for Intelligent Systems*, pp. 57–61. ACM.
- Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Harvard University Press.
- Wermter, S. a M. Elshaw (2003). Learning robot actions based on self-organising language memory. *Neural Networks* 16(5-6), pp. 691–699.